

児童における算数学習の自己評価と内発的動機づけとの関連[†]

鈴木 教益*・内野康人之**

石橋町立石橋小学校*

宇都宮大学教育学部**

今日の教育実践上の課題として子どもの学習意欲の向上に有効な自己評価のあり方が求められている。本論文での主たるテーマは、到達度評価による自己評価が小学校5年生の算数学習と動機づけに及ぼす効果を実験授業によって検討した鈴木（2004）において解明が十分でなかった問題について、データの再分析と新たなデータの追加によって考察を深めることにある。その結果、授業レベルでの自己評価において授業目標を意識した分析的評価を媒介とする総合的評価の遂行が、少数ではあるものの動機づけ水準が低く授業での習得水準が十分でない児童の動機づけを高める可能性を確認した。また、ハーター（1981）による内発的・外発的動機づけ測定尺度に見られる児童の1年間の変化は特に「基準」尺度に顕著であった。到達度評価が教師と子どもによる評価基準の共有化を図る条件を有していることに鑑みて、自己評価による学習意欲向上の出発点となる手立てがまず評価基準の学び（内面化）に求められることが示唆された。

キーワード：小学生、算数学習、自己評価、到達度評価、内発的動機づけ

はじめに

学習への動機づけ（学習意欲）は学習活動を喚起し支え持続させるエネルギーである。それは学習活動の過程と成果にとって重要な条件であると同時に、学習活動による効果でもある。したがって、子どもにおける「確実な学習」と「発達の充実」を支援する学校教育では、学習意欲の向上を図ることが重要な実践的課題となる。そのために、教師は子どもの興味・関心を引くような課題や教材を工夫したり、学級での人間関係に配慮したりする。しかし、教師による評価も子どものインフォーマルな自己評価を介して学習意欲に影響を及ぼしていることは見

落とされがちである。

1989（平成元）年改訂の学習指導要領において「自ら学ぶ意欲」の育成が重視され、評価の面でも「指導の過程や成果を評価し、指導の改善を行うとともに、学習意欲の向上に生かすように努めること」とされ、自己評価に学習意欲向上の手立てとなることが期待された。さらに、学習指導要領改訂と連動した1991年の指導要録改訂では、評価の指導機能が強化されると同時に「観点別学習状況の記録」において、意欲とそれによって統合される関心・態度が4観点の冒頭に位置づけられた。しかし意欲などの情意的側面の評価は困難である。そこで、学習意欲を評価する手段として子どもによる自己評価が導入されもした。

このような教育状況のなかで、教育心理学の

[†] Takayoshi SUZUKI* and Yasutoshi UCHINO**:
Self-evaluation on Arithmetic Study and Intrinsic
Motivation in Elementary School Children.

* Isibashi Elementary School

** Faculty of Education, Utsunomiya University

分野でも評価構造が内発的動機づけに及ぼす効果をテーマとする研究が目立つようになった。

内発的動機づけの概念については曖昧な面があるが、行為の手段性－目的性の区別に基づく概念化が広く受け入れられている（鹿毛,1994¹⁾）。例えば、他者からの物質的報酬を求めたり、叱責を避けたりするための手段として学習に取り組んでいる場合を外発的動機づけと言い、課題に興味があるから、疑問に思ったから学習そのものを目的として自発的に取り組んでいる場合を内発的動機づけと言う。ハーター（1981）は、内発的動機づけを教室における学習や熟達への指向性として捉え、その個人差を測定する尺度を開発した（鹿毛,1994）。具体的には「挑戦」「知的好奇心」「達成」という内発的動機づけの中身とも言うべき3つの下位尺度と、内発的動機づけに影響を与える認知的評価能力に関する「判断」「基準」の2つの下位尺度から成る（桜井,1990²⁾）。

算数・数学学習において授業終了期に学習状況（理解や技能）に関する小テストを実施し、それを自己評価させるという手続きを組み込んだ実験授業の効果を内発的動機づけの測定によって検証するといった枠組みでの研究に焦点を当ててみると、鹿毛（1993³⁾）、西松・千原（1995⁴⁾）、中川・松原（1996⁵⁾）の研究がある。

子どもの場合には、とにかく自己評価をさせれば、それがストレートに学習意欲を高めるとは思われない。自己評価活動は自己が評価の主体であると同時に評価の対象でもあるという側面によって特徴づけられるが、活動であるからには目的、対象の範囲、時期、方法、情報等の要素も大切である。これらは学校教育では教師の支援に属する。そこで、これらの教育心理学的研究から、内発的動機づけの向上に着実に結びつく自己評価の条件を探っていくことが必要だと思われる。

鹿毛（1993）は、小テストの「正答数」を学習目標と対応づけて「目標を達成したかどうか」の情報を自分で生み出す自己評価活動が小学5年生の分数量学習への「知的好奇心」「自律性」及び「挑戦」を向上させることを明らかにした。到達度評価情報が学習の目標・内容と直接的に関係しているがゆえに、問題が単に「できた・できなかった」から内容が「わかった・わからない」へ注意を向けさせると同時に有能感を高め、それを媒介として内発的動機づけを高めていると考えられる。また、テスト活動後に自分で採点し評価情報を生み出す活動は、評価過程での積極的にかかわりを求め、自己決定感を感知させるため、内発的動機づけを高める可能性をもつと考えられる。そして、到達度評価情報を生成するためには自己採点から自己評価への手立て（例えば評価基準の学び）が必要であろう。

中川・松原（1996）は、割り算の学習において自己評価の対象を学習成果だけでなく、授業の目標に対応した解決方法にまで広げることが、小学3年生の内発的動機づけに肯定的影響を及ぼしていることを明らかにした。すなわち、ハーター（1981）による測定尺度の下位尺度の面で「基準」や「達成」の上昇的変化が観察されたのである。

西松・千原（1995）は、個人内評価に基づく教師評価と子どもの自己評価が同時に実施されたときに、中学1年生の正の数・負の数の学習への内発的動機づけを強く喚起することを明らかにした。ここでの自己評価は「理解の程度」や「復習の必要性」についてであるので、到達度情報とは言い切れないが、それでも効果が見られたのは教師からの個人内評価（自己の学習活動のよくできた点・できていない点及び前回からの進歩の様子）を得たからであろう。学校教育では教師からの評価が子どもにとって大き

な意味をもつことを前提として（教師からの評価は単なる外からの評価ではないし、とりわけ到達度評価情報のフィードバックは教授＝学習活動に内在的なものである）、教師評価の支えのなかでの自己評価の条件を探ることが必要であると思われる。

総括的テストを指標とした学習効果が見られたのは中川・松原（1996）においてである。学習方法の自己評価を含んでいたからであろう。小テスト得点に関しては鹿毛（1993）において自己評価の効果が見られた。したがって、評価→動機づけ→習得→評価→動機づけのサイクルのなかで、子どもによる自己評価を含めて「習得と意欲を媒介する評価の機能」に注意を向ける必要があろう（内野，1995⁶⁾）。

1998(平成 10)年に改訂された学習指導要領では「自ら学び自ら考える力の育成」を重視している。引き続き指導要録の改訂（2002）では学習評価において目標に準拠した評価（到達度評価）を「評定」でも採ることになった。学習と意欲をつなげる評価の機能や方法に関する研究的実践にとって好条件がもたらされたと考えたい。

鈴木（2004⁷⁾）は、2003 年 5 月、栃木県内 I 小学校の 5 年生を対象に到達度評価を軸にした多面的な自己評価を取り入れた実験授業（実験群）が学習と内発的動機づけに及ぼす効果をハーター（1981）による内発的－外発的動機づけ測定尺度によって検討した。統制群との比較において総括的テストの一部に学習効果が見られたものの、動機づけへの効果は十分に確認できなかった。そこで本論文では、この問題について、一方で分析視点を広げ、他方で新たなデータを追加することによって考察を深めることを主たる目的とする。

新たなデータとは、鈴木（2004）の実験群となった学級において 2004 年 3 月 16 日に、

算数授業の一環として実施した算数の計算を中心にした自己評価、及びその 2 日後に実施した内発的－外発的動機づけ測定尺度である。

テーマ 1. 内発的－外発的動機づけ測定尺度に見る動機づけの変化

鈴木（2004）での実験群となった学級（38 人）においてハーターによる内発的－外発的動機づけ測定尺度（以下、単に「動機づけ測定尺度」と呼ぶ）を 3 回実施した。実施時期は 1 回目が 2003 年 5 月 6 日、2 回目がほぼ 1 カ月後の 5 月 30 日、そして 3 回目が学年末の 2004 年 3 月 18 日であった。

動機づけ測定尺度（各 6 項目の 5 下位尺度から成る）は桜井・高野（1985⁸⁾）の翻訳によるものを用いたが、原尺度中「判断」に属する 2 項目を児童の実態や学校の教育実践に照らして実情にそぐわないと考え、改変した。鈴木（2004）における学年当初の 5 年生 105 人（実験群を含む）のデータの因子分析結果によると、改変項目は、他の判断項目（原尺度）が属している因子の負荷量が相対的に高かったので、概ね妥当性ありと判断される。

回答は外発的動機づけを示す内容と内発的動機づけを示す内容の対のいずれかに対して「だいたいあてはまる」か「よくあてはまる」かを求めるもので、構造としては 4 段階評定である。得点は内発的動機づけを示す内容に「よくあてはまる」を 4 点とし、外発的動機づけを示す内容に「よくあてはまる」を 1 点としたので、得点範囲は尺度全体で 30 ～ 120、下位尺度で 6 ～ 24 となる。表 1 に 3 回の結果を示す。

（1）1 回目と 3 回目の比較

両者の相関係数は一般的には尺度全体の信頼性係数として見ることになるだろうが、信頼性係数としての数値はそれほど高くないので、多様な変化を想定した方がよい。対応のある t 検定に

よって、ほぼ1年間の差異を検討したところ、尺度全体で得点の上昇（内発的動機づけの方向への発達）が見られた。これを別の指標によって補足するならば、項目当たり3点（あてはまる）を超える91点以上（総合的に見て内発的動機づけの傾向にある）の子どもの人数が19人（50%）から27人（71%）に増えている。

発達は「基準」尺度において顕著であった。小学3年から中学3年までの子どもについて横断的方法によって発達傾向を調べた桜井（1990）は、基準に関して外発的傾向から内発的傾向へと転換し始める時期は小学5年あたりからであり、その後も発達していくという結果を示しているが、本結果も、このような発達傾向のプロセスのなかで解釈できよう。

進級に伴うクラス替えによる緊張と探索がおさまり比較的落ち着いた時期での1回目の結果は、学年初めのベースラインと見ることができる。1回目の水準が91点以上の者（19人；平均98.84，標準偏差7.48）と90点以下の者（19人；82.58，6.90）の2群に分けて、各群における変化を比較すると、前者では100.37（12.59）に上昇したものの有意な差ではなかった。対して後者では、88.63（12.24）に有意に上昇した（ $t=2.317$ ， $p<.05$ ）。平均で見る限り「内発的動機づけの傾向にある」とは言い切れないが、19

人中11人が91点以上の水準への変化であることに注目したい。また、この群で「基準」の他に「挑戦」尺度にも5%水準で有意な上昇的变化が見られたのに対して、91点以上の群では「基準」のみに有意な傾向が見られた。

挑戦と基準は比較的明瞭なまとまりをもった別個の因子を構成している（桜井，1990；鈴木，2004）ので、内発的動機づけの発達を捉えるのに適していると考えられるが、挑戦が先行して発達するのではないかと思われる。

（2）1回目と2回目の比較

鈴木（2004）では、この比較は多面的な自己評価を積極的に取り入れた実験授業の効果を検討するために行われた。

表1に見るように、尺度全体では変化に有意差はなく、相関係数もかなり高いことから約1カ月経過後の尺度の信頼性を確認したかのである。しかし、1回目と3回目の比較では検出されなかった「知的好奇心」の下降と「挑戦」の上昇が見られたことは、それらが一時的であるにしても実験条件によるものとみなした上での検討が必要であろう。また、「基準」の上昇についても発達傾向のプロセスを示すという見方をする前に、実験条件の効果とみなす根拠を探ってみる価値がある。これが、テーマ2での検討課題である。

表1 内発的－外発的動機づけ測定尺度得点の推移，調査時期間の相関係数及び変化の検定

		1回目	2回目	3回目	2－1回目 相関 t		3－1回目 相関 t	
全 体		90.71(10.87)	90.76(12.44)	94.50(13.61)	.72	.036	.59	2.058*
下 位 尺 度	挑 戦	17.47(4.08)	18.66(4.42)	18.53(4.67)	.63	1.998+	.49	1.463
	達 成	20.18(2.59)	19.71(2.94)	20.03(3.20)	.55	-1.107	.36	-.293
	知的好奇心	16.82(2.89)	16.03(2.88)	17.39(3.06)	.68	-2.092*	.59	1.323
	判 断	20.29(2.65)	19.63(2.42)	20.37(2.59)	.26	-.658	.47	.181
	基 準	15.95(4.22)	16.74(4.85)	18.18(4.29)	.86	1.968+	.41	2.979**

数値は平均，（ ）内は標準偏差，+ は $p<0.10$ ，* は $p<0.05$ ，** は $p<0.01$

テーマ2. 実験条件としての自己評価と動機づけとの関連

鈴木（2004）の実験授業についてもう少し詳しく紹介しておこう。16 時間（1 単位時間は 45 分）扱いの算数単元「小数と整数のかけ算，わり算」の指導において多面的な自己評価を実施した。特別の目標，教材，指導方法，学習時間，ドリルプリント，テスト等を用いたのではなく，統制群の学級と同様に“ふつう”の授業を行ったが，単元全体を通して①授業レベルから小単元レベルへ，②教師が用意した評価観点から子どもが意識した評価観点へ，③テスト問題の正誤のレベルから目標の達成状況のレベルへ，④個別的・分析的評価から全体的・総合的評価へといった多面的な自己評価を取り入れた評価実践を試みた。そして評価実践全体をセットにして統制群法によってその効果を検討したのである。有効と思われる要因をつなげてそれらの相乗効果をねらうのが教育実践の筋だとすれば，このアプローチは間違っていないが，統制群との比較に関する限り，動機づけへの効果は確認できなかった。

そこで，実験群のみを対象として，授業レベルでの分析的評価から総合的評価への自己評価の遂行が動機づけ測定尺度における個人内変化にどのような影響を及ぼしているかを検討した。その自己評価の概要は以下のようである。

1 時間の授業の終期に「ドリルプリント」（文溪堂）を実施し，子どもによる自己採点后，教師が用意した「自己評価カード」によって自己評価をさせた。自己評価カードは，中川・松原（1996）を参考にして，①正答数を記入する「できた数」，②計算の仕方に関する当該授業の目標（3つ）に照らして目標ごとに「よくできた」から「もう少し」までの5段階評定を求める「みつめてみよう」，③当該授業での学習状況を総合的に5段階で評価させる「できたか

な」の3つの内容から成る。この自己評価は，小単元「小数と整数のかけ算」（4時間）で2回，小単元「小数と整数のわり算」（8時間）で5回の，計7回実施された。

分析の視点は「できたかな」の総合的評価が何に基づいているかに拠った。少なくとも「できた数」と「みつめてみよう」の2つが考えられるが，授業目標に注意を向けて総合的評価を行う態度を育てたいとの思いからすれば，「みつめてみよう」の分析的評価を媒介とした総合的評価が望ましいし，またその評価方法は教師が子どもに自己評価させる場合に取り得る手立てになると考えたからである。

そこで「自己評価カード」の記録から，7回の自己評価において，

＜A群＞総合的評価が正答数よりも分析的評価に対応している者（11人），

＜B群＞総合的評価が分析的評価よりも正答数に対応している者（6人），

＜C群＞正答数も分析的評価も高く，総合的評価がいずれによるかを識別できない者（14人）

＜D群＞総合的評価が何に基づいているかが不明，または対応関係が一貫していない者（7人），の4群に分けた。

動機づけ測定尺度において5点以上（下位尺度当たり1点以上）の上昇的变化が見られた者の比率がA群で最も高かった（11人中5人；46%）ので，一応，授業目標を意識した分析的評価を媒介とする総合的評価が有効でありそうだと結論を得た。しかし，用いた測度に現れない子どもの変化を見取ることのできる実践者の目を踏まえているとはいえ，データの上で十分に根拠があるとはいえない。そこで，この結論を作業仮説とした上で，さらに別の視点からの検討を踏まえながら検証を試みる。これがテーマ2の課題の1つである。

(1) 動機づけ測定尺度における個人内変化を 従属変数として

動機づけ測定尺度は実験授業開始の前日及びその終了の翌日に実施した。

主に「自己評価カード」の記録に基づき4つの視点から児童を分類し、「知的好奇心」「挑戦」及び「基準」の変化がどの群に見られるかを検討する。

取り上げる視点について説明した上で、対応のあるt検定の結果(表2)について述べることにする。

【視点1】は鈴木(2004)における上記の視点である。

どの群においても尺度全体での変化は有意でなかった。下位尺度では、B群に基準の上昇が、C群に挑戦の上昇が見られた。

【視点2】は「できたかな」の総合的評価のみに着目しての推移のパターンである。7回の自己評価における「よくできた」(5段階の「5」)の比率は、65.8%、65.8%、73.7%、55.3%、42.1%、68.4%、78.9%のように推移している。もちろん多くの子どもにおいて、この比率は正答数(習得の実態)に対応している。ちなみに、8～11問から成るドリルプリントにおいて満点者の比率の推移は、68.4%、68.4%、84.2%、44.7%、42.1%、68.4%、84.2%のようであった。小数と整数のわり算での困難性に応じて、4回目(商が1より小さくなる計算、除数が2ケタ)と5回目(わり進むわり算、わりきれないわり算)の学習内容での誤答が目立つこの時期では「5」の減少を始め、多くの子どもの総合的評価が低下している。

そこで、「5」に関して、①多い(4回以上)か少ない(3回以下)か、②7回目が「5」の評価で終わるか終わらないか、の組み合わせを基本に以下のように分類した。

<A群>すべての総合的評価において「5」

の者(11人)、

<B群>「5」が多く、7回目が「5」で終わっている者(13人)、

<C群>「5」が少なく、7回目が「5」で終わっている者(6人)、

<D群>「5」が多く、7回目が「5」で終わっていない者(1人)、

<E群>「5」が少なく、7回目が「5」で終わっていない者(7人)、の5群に分けた。

B群において挑戦と基準の上昇が、E群に知的好奇心の下降が見られた。

【視点3】は7回の「できた数」を通して誤答が少ないか多いかという視点である。これは子どもの自己評価というよりも、習得の実態に対する教師サイドからの評価に通じる視点である。満点の回が、5回以上を<A群>(21人)、4回以下を<B群>(17人)とした。

A群では挑戦と基準の上昇が、B群では尺度全体の下降とともに知的好奇心と達成の下降が見られた。

【視点4】は1回目と3回目との比較でも取り上げた1回目の動機づけ水準による分類である。91点以上を<A群>(19人)、90点以下を<B群>(19人)とした。この視点も視点3以上に、子どもの自己評価とは直接的な関係がない。

B群において基準の上昇が見られた。

以上の結果を総合的に考察するに当たっては、視点間での群の重なりには留意する必要があるが、その記述は煩雑になるので、最小限に止める。以下、主要な考察について述べる。

①尺度全体における有意な変化は、子どもの自己評価に基づいていない視点3のB群にのみ見られたので、この視点が動機づけの変化に敏感なのかもしれない。授業での習得水準が動機づけの変化の規定要因であり、学習中に誤りが多いと動機づけを低下させるという一般的知見

表2 各分析視点における群間の比較

	動機づけ測定尺度得点 (群内比較)				情意得点 (群間比較)			
	群 尺度	1回目平均	2回目平均	t	群-群	前群の平均	後群の平均	t
視点1	B 基準	15.33(5.64)	17.67(6.92)	2.445+	C-A	11.29(0.73)	8.73(2.41)	3.776**
	C 挑戦	19.64(3.37)	21.00(3.62)	1.857+	C-B	11.29(0.73)	9.67(1.51)	3.301**
					C-D	11.29(0.73)	8.86(1.95)	4.195***
視点2	B 挑戦	17.46(3.60)	20.00(3.49)	2.008+	A-E	10.73(2.05)	7.86(2.54)	2.638*
	基準	17.62(3.48)	18.92(3.43)	2.168+	B-E	10.31(1.38)	7.86(2.54)	2.825*
	E 好奇	16.71(3.25)	14.86(2.34)	-1.982+				
視点3	A 挑戦	18.29(3.95)	20.33(3.89)	2.387*	A-B	10.57(1.29)	8.94(2.38)	2.691*
	基準	17.00(3.49)	18.00(4.31)	1.811+				
	B 全体	89.06(11.86)	85.59(13.03)	-1.792+				
	達成	20.24(2.97)	18.59(3.54)	-2.865*				
視点4	好奇	17.06(2.66)	15.76(2.75)	-2.335*				
	B 基準	13.68(3.79)	15.05(4.93)	2.387*	A-B	10.53(1.68)	9.16(2.12)	2.209*

数値は平均, ()内は標準偏差, +は $p<0.10$, *は $p<0.05$, **は $p<0.01$, ***は $p<0.001$

を確認したことになる。したがって、授業での学習を確実にするための指導が動機づけを低下させないための最良の方策である。

②視点3と4には群間の重なりがほとんどない(動機づけ水準の高い者が習得水準の高い群に集中しているわけではない)ので、両者を関連させて分析した結果、動機づけ水準が低くても授業での習得水準が高ければ「挑戦」の上昇が、適性要因としての動機づけが高くても習得が低ければ「達成」の下降が予想される。

③視点4のB群における「基準」の上昇は、動機づけ水準が高くない子どもに固有の変化であるようだ。鹿毛(1993)は到達度評価による自己評価が外発的動機づけの傾向にある子どもの小テストへの肯定的態度と自律性の向上に寄与しているという結果を得ているが、本結果でも頻繁な小テストの自己評価を経験するなかで、できた・できないの評価基準について教師からの基準への依存が次第に少なくなったと考えられる。評価の面から動機づけの向上を図る

方策の出発点に評価基準の学び・内面化が位置づけられることが示唆される。

④視点3と2では群間に重なりがある。この背景に、正答数(習得水準)と総合的評価との関連の大きさがある。これを敷衍すれば正答数に基づく教師評価と子どもの自己評価との関連になる。時には両者にズレがある。そこに自己評価の独自な意味があるが、こと学習評価に関しては一致が望ましい。評価基準の押し付けではなく共有によって一致を図るべきである。

到達度評価では、満足できる習得水準についての基準が必要である。例えば「観点別学習状況」の各観点において「おおむね満足できる」の習得基準は教師が設定するにしても、子どもに受け容れられてなくてはならない。子どもはそれを内面化し自己評価の際に利用する。この評価基準の学び(内面化)を支えるのは教師の役割であろう。鈴木(2004)では、この点において不十分であった。

誤答が最多であった5回目のドリルプリント

において正答数と総合的評価のズレが見られた 8 人中 5 人は視点 1 の A 群であった。

⑤授業における習得水準に対する総合的評価の推移パターンに基づく視点 2 では、一まとまりの学習活動の途中で評価の低下がある B 群の結果は、それが無かった A 群との比較において意味がある。多少困難な課題に直面して誤りが生じたとき、評価を落としながらもどこに問題があるかの診断的情報を得て、努力を強め、そしてそれが報われるならば成就感や有能感を感知して、さらに困難な課題に挑戦しようとするだろう。新しい内容の学習に誤りは付き物であるので、このような形成的評価の機能をもつ自己評価が望ましいと考えられる。

⑥視点 2 の B 群では、「挑戦」と「基準」の相関係数が 1 回目での .018 から 2 回目で .544 へ変化している。桜井（1990）では .30 ～ .35 の数値が得られている。ちなみに 38 人のデータでの 1 回目から 3 回目までの推移は、.330, .394, .542 のようであるので、この 2 つの下位尺度間の関連が緊密になることが内発的動機づけの発達における重要な変化ではないかと推察される。

⑦鈴木（2004）の視点 1 に立ち戻ったとき、A 群は視点 2 で B 群と E 群に、視点 3 で A 群と B 群に分散しているところに着目したい。尺度全体で 5 点以上の上昇的变化が見られた子ども 12 人中 5 人が視点 1 の A 群であり、その中の 3 人は視点 2 - 3 において E 群 - B 群のパターンであった。したがってこれまでの分析によれば動機づけの低下が予想されるのに、そうではなかった。尺度全体のみならず「基準」の上昇も見られたが、3 人とも 5 回目のドリルプリントでのズレは見られなかった。また、1 回目の動機づけ水準は 3 人とも 75 点以下であったが、後述する小単元レベルでのテスト結果に基づく自己評価の後での感想は「がんばって楽し

かった」「算数がこの勉強で少し好きになった」「全体的にできている」というように学習活動に対して肯定的なものであった。

分析的評価を媒介とした総合的評価は、そのような仕掛けを自己評価カードに作ったわけであるので、記録からは識別できない視点 1 の C 群の子どもにもインフォーマルに生じていたかもしれない。したがって、このような自己評価のあり方を推奨するような教師からの働きかけが必要でもあった。また、その動機づけへの効果を確認するには適切な検証の方法が求められる。しかし、少数ではあるが動機づけの向上が見られたことは、分析的評価を媒介とした総合的評価をフォーマルに意図的に行わせるような評価実践であれば、もっと多くの子どもに対して有効であろうと推測される。

⑧実験群全体に、そして視点 2 の E 群及びその群のほとんどが含まれている視点 3 の B 群に、「知的好奇心」の下降が見られた。これは一見、問題点として映る。

桜井（1990）によれば、知的好奇心 6 項目中「よい成績を取るために一生懸命勉強する」と及び「よい点を取るために一生懸命勉強する」という 2 つの外発的動機づけを示す内容には多くの子どもが「あてはまる」と回答している。また、それらは他の知的好奇心項目とは別の因子を構成しており、その発達傾向としては小学 5 年あたりから下降している。鈴木（2004）の因子分析結果でも、知的好奇心 6 項目は 5 因子中 3 つの因子に分かれており、先の 2 項目は、「先生に正しい答えを教えてもらわなくても、自分の答えが正しいかどうか分かる」と及び「テストの点を見なくても、よくできたかどうか予想できる」の基準尺度項目と同じ因子に高い負荷をもっている。「基準」の上昇が「知的好奇心」の下降とリンクしていることを示唆しているが、実験条件のなかでも総括的テスト結果に

基づく自己評価がこのリンクを強化した可能性がある。小単位における目標全体の達成状況をテスト結果に基づく個々の目標達成状況から総合的に評価するときに「よくできた(A)」と自己評価した者に、この2つの項目での低下が目立った。例えば「よい成績を取るために一生懸命勉強する」では28人中9人に低下が見られたのに対して、B及びCの評価では合わせて10人中1人であった。テスト状況での自己評価がよい成績や点数に注意を向かわせ、その方向での動機づけ(外発的動機づけ)を強めることに留意すべきであろう。

実験群全体に見られた知的好奇心の下降のある程度は、このケースであった。

⑨ハーターによる動機づけ測定尺度は、特性レベルでの一般的な動機づけの個人差を把握するためのものである。学習内容に密着した動機づけの変化があったとしてもそれをすくい上げることが難しいという問題点がある。

これが、鈴木(2004)において、統制群との比較でも個人内変化でも動機づけへの効果を十分に確認できなかった一因であるかもしれない。

(2) 小単位レベルでの情意的側面の自己評価を従属変数として

ここでの検討は、鈴木(2004)で未分析のものである。

小単位の指導終了後に総括的テスト(青葉出版社)を実施し、教師採点のテストを翌日に返却した上で、教師が用意した評価表での自己評価を求めた。それは、次の手順で進められた。

①テスト結果から小単位での目標の達成状況を目標ごとに評価する。

②目標全体の達成状況を個々の目標達成状況から総合的に評価する。

③学習活動における情意的側面について実現状況を評価する。

④情意的側面全体の実現状況を個々の側面の実現状況から総合的に評価する。

⑤目標全体の達成状況と情意的側面全体の実現状況とに基づき小単位の学習状況を総合的に評価する。

いずれも回答は「よくできた(A)」「ふつう(B)」「あまりよくできなかった(C)」の3段階評価による。

③の学習活動における情意的側面についての振り返りのために用意した項目は、「楽しく学習できた」「自分から進んで学習できた」「自分なりの考えをもって学習できた」及び「むずかしい問題でもあきらめずにとりくんだ」の4項目である。これらは普段の授業で子どもの情意的側面における状況を教師が理解するためによく用いているものである。

3段階評価による回答の4項目全体(「情意得点」と呼ぶ)の得点範囲は4～12点であり、平均が9.92、標準偏差が1.81であった。

この測度は授業レベルでの自己評価の従属変数となり得るので、(1)の4つの各視点においてt検定による群間比較を行った。

表2に見るように、情意得点における有意差は、視点1ではC群がいずれの群よりも高く、視点2ではA群とB群がともにE群よりも高く、視点3ではA群がB群よりも高いところに見られた。したがって、情意得点は学習状況を一定程度反映している。他方、視点4においてA群がB群よりも情意得点が高いことは、実験授業前の動機づけ水準も反映されていると思われる。ちなみに、1回目の動機づけ測定尺度の全体得点と情意得点との相関係数は.451であり、下位尺度では、項目内容との関係が濃いと思われる「挑戦」と.432の相関が見られたのに対して、「基準」とは.238と低かった。したがって、ふつう用いられている項目では、子どもに生じる可能性の高い動機づけの側面での変

化を見落としてしまうことにもなりかねない。

(3) 学年末での自己評価からの検討—ドリル プリントの自己評価の経験—

学年末での自己評価の中心は、①計算学習における理解度の自己評価、②計算への動機づけ、の2つに置いた。紙幅の都合上、これら両者の内容及び関係を中心とする検討については別稿に譲ることにして、ここでは、これまでの検討を補足する1つの質問項目に対する結果のみを報告する。

鈴木(2004)の実験群では、5月(小数と整数の乗除算の学習)の他に11月(同分母の分数の加減算の学習)にもドリルプリントでの自己評価を行っている。当時の「自己評価カード」のファイルを返却することによって記憶を確かにしながら、それらの経験を振り返っての気持ちを問うたところ、「楽しくできた」への回答が29人(76%)に見られた。しかも、1回目の動機づけ水準が高くない(90点以下)の者19人中17人が楽しいと回答した($\chi^2=3.640$, $p<0.10$)、2回目のデータに基づいても同様であった(18人中16人; $\chi^2=2.400$, $p<0.10$)。他方、3回目のデータとのクロスでは有意差がなかった。これらの結果は、テーマ2で検討した視点4のB群における授業レベルでの自己評価の効果を傍証するとともに、テーマ1で述べた内発的動機づけの発達の源泉の一部が自己評価を導入した評価実践にもあったことを示唆するものである。

逆に、「役に立った」(23人;61%)に関しては、3回目の動機づけ水準とのみ関連があった($\chi^2=3.783$, $p<0.10$)。91点以上の者に、この回答が多く見られた(27人中19人)が、1年間の変化に有意差はなかった。

最後に、「授業で何が大切か考えるようになった」に16人(42%)、「他の勉強でも生かしている」に15人(39%)、「その後も自分で自

己評価している」に7人(18%)の回答があったが、いずれも動機づけ水準とは無関係であり、それらの回答者を特徴づける動機づけでの変化も見られなかった。しかし、「他の勉強でも生かしている」への回答者で、計算への動機づけにおける「基準」得点が高かった($t=2.147$, $p<0.05$)ことは、ハーターの動機づけ測定尺度では拾えない動機づけの側面があることを暗示している。

参考文献

- 1) 鹿毛雅治(1994) 内発的動機づけ研究の展望 教育心理学研究,42, 345-359.
- 2) 桜井茂男(1990) 内発的動機づけのメカニズム 風間書房.
- 3) 鹿毛雅治(1993) 到達度評価が児童の内発的動機づけに及ぼす効果 教育心理学研究,41, 367-377.
- 4) 西松秀樹・千原孝司(1995) 教師による個人内評価と自己評価が生徒の内発的動機づけに及ぼす効果 教育心理学研究, 43,438-444.
- 5) 中川恵正・松原千代子(1996) 児童における「わり算」の学習に及ぼす自己評価訓練の効果—自己評価カード導入の効果— 教育心理学研究,44, 214-222.
- 6) 内野康人之(1995) 教育評価と動機づけ(新井邦二郎 編著 教室の動機づけの理論と実践 8章) 金子書房.
- 7) 鈴木教益(2004) 到達度評価による自己評価が児童の学習及び内発的動機づけに及ぼす効果 宇都宮大学教育学研究科修士論文(未公刊).
- 8) 桜井茂男・高野清純(1985) 内発的—外発的動機づけ測定尺度の開発 筑波大学心理学研究,7, 43-54.